

METHOD AND APPARATUS FOR DATA PROCESSINGTechnical Field

5 The present invention relates to the technical field of computer-aided information management, and concerns more specifically a method and an apparatus for data processing ~~according to the preamble to claim 1 and claim 8~~ ^{respectively} for accomplishing increased protection against ~~unauthorised~~ ^{unauthorized} processing of data.

Background Art

10 In the field of computer-aided information management, it is strongly required that the protection against ~~unauthorised~~ ^{unauthorized} access of data registers be increased, especially against violation of the individual's personal integrity when setting up and keeping personal registers, 15 i.e. registers containing information on individuals. In particular, there are regulations restricting and prohibiting the linking and matching of personal registers. Also in other fields, such as industry, ~~defence~~ ^{insurance}, banking, insurance, etc, improved protection is desired against ~~unauthorised~~ ^{unauthorized} access to the tools, databases, applications 20 etc. that are used for administration and storing of sensitive information.

WO95/15628, which has the same owner as the present application, discloses a method for storing data, which 25 results in increased possibilities of linking and matching with no risk of reduced integrity. The method, which is illustrated schematically in Figs 1 and 2 on the enclosed drawing sheets, concerns storing of information comprising on the one hand an identifying piece of information or original identity OID, for instance personal 30 code numbers Pcn and, on the other hand, descriptive information DI. The information OID + DI is stored as records P in a database O-DB according to the following principle:

- Step 1 OID (Pcn) is encrypted by means of a first, preferably non-reversible algorithm ALG1 to an update identity UID;
- Step 2 UID is encrypted by means of a second, reversible algorithm ALG2 to a storage identity SID;
- Step 3 SID and DI are stored as a record P in the database O-DB, SID serving as a record identifier;
- Step 4 At predetermined times, an alteration of SID in all or selected records P is accomplished by SID of these records being decrypted by means of a decrypting algorithm ALG3 to UID, whereupon UID is encrypted by means of a modified second, reversible algorithm or ALG2' to a new storage identity SID', which is introduced as a new record identifier in the associated record P as replacement for previous SID. This results in a security-enhancing "floating" alteration of SID of the records.

For a closer description of the details and advantages of this encrypting and storing method, reference is made to WO95/15628, which is to be considered to constitute part of the present description. The storing principle according to steps 1-4 above is ^{herein} ~~below~~ referred to as PTY, which is an abbreviation of the ^{principal} ~~concept~~ of PROTEGRITY which stands for "Protection and Integrity".

A detailed technical description of PTY is also supplied in the document "PROTEGRITY (ASIS) Study 2", Ver. 1.2, 1 March 1996, by Leif Jonson. Also this document is to be considered to constitute part of the present description.

In the technical field at issue, so-called shell protections, ~~however~~, are today the predominant method of protection. Shell protection comprises on the one hand the external security (premises) and, on the other hand, an ^{authorization} ~~authorisation~~ check system ACS with user's passwords for controlling the access. ACS is used as shell protection for main frames, client/server systems and PC,

but it does not give full protection and the information at issue can often relatively easily be subjected to ~~unauthorised~~ ^{unauthorized} access. This protection has been found more and more unsatisfactory since, to an increasing extent, 5 "sensitive" information is being stored, which must permit managing via distribution, storing and processing in dynamically changing environments, especially local distribution to personal computers. Concurrently with this development, the limits of the system will be more and 10 more indistinct and the effect afforded by a shell protection deteriorates.

Summary of the Invention

In view of that stated above, the object of the present invention is to provide an improved method for processing information, by means of which it is possible to increase the protection against ~~unauthorised~~ ^{unauthorized} access to sensitive information.

A special object of the invention is to provide a 20 technique for data processing or managing, which makes it possible for the person responsible for the system, the management of the ~~organisation~~ ^{organization} etc. to easily establish and continuously adapt the user's possibility of processing stored information that is to be protected.

A further object of the invention is to provide a 25 technique for data processing which offers protection against attempts at ~~unauthorised~~ ^{unauthorized} data processing by means of non-accepted software.

One more object of the invention is to provide a 30 technique for data processing according to the above-mentioned objects, which can be used in combination with the above-described PTY principle, for providing a safety system with an extremely high level of protection.

~~These and other objects of the invention are achieved by the method according to claim 1 and the apparatus according to claim 8, preferred embodiments of the invention being stated in the dependent claims,~~ 35

Thus, the invention provides a method for processing of data that is to be protected, comprising the measure of storing the data as encrypted data element values of records in a first database (O-DB), each data element
 5 value being linked to a corresponding data element type.

The inventive method is characterised by the following further measures:

storing in a second database (IAM-DB) a data element protection catalogue, which for each individual data
 10 element type contains one or more protection attributes stating processing rules for data element values, which in the first database are linked to the individual data element type,

in each user-initiated measure aiming at processing
 15 of a given data element value in the first database, initially producing a compelling calling to the data element protection catalogue for collecting the protection attribute/attributes associated with the corresponding data element type, and compellingly controlling the processing
 20 of the given data element value in conformity with the collected protection attribute/attributes.

In the present application the following definitions are used:

- 25 ° "Processing" may include all kinds of measures which mean any form of reading, printing, altering, coding, moving, copying etc. of data that is to be protected by the inventive method.
- C Insc
 C ° "Data element type" ^{identifies} ~~concerns~~ a specific ^{category} ~~type~~ of data ^C
~~having a meaning as agreed on.~~
- 30 ° "Data element value" concerns a value which in a given record specifies a data element type.
- 35 ° "Record" concerns a number of data element values which belong together and which are linked to the respective data element types, optionally also including a record identifier, by means of which the record can be identified. Example:

70060

DATA ELEMENT TYPE		
RECORD ID	SOCIAL ALLOWANCE	CAR
XXXX XXXXX	encrypted data element value	encrypted data element value
YYYY YYYY	encrypted data element value	encrypted data element value

- "Protection attribute indicating rules of processing" may concern:

- data stored in the data element protection catalogue and providing complete information on the rule or rules applying to the processing of the corresponding data element, and/or
- data stored in the data element protection catalogue and requiring additional callings to information stored in some other place, which, optionally in combination with the protection attributes, states the processing rules involved.

- "Collection of protection attributes" may concern:

- collection of the protection attributes in the form as stored in the data element protection catalogue, and/or
- collection of data recovered from the protection attributes, for instance by decryption thereof.

- "Encryption" may concern any form of encryption, tri-cryption, conversion of coding of plain-text data to non-interpretable (encrypted) data, and is especially to concern also methods of conversion including hashing.

The inventive method offers a new type of protection, which differs essentially from the prior-art shell protection and which works on the cell or data element level. Each data element type used in the records in the first database is thus associated with one or more protection attributes, which are stored in a separate data element protection catalogue and which protection attri-

6

butes state rules of how to process the corresponding data element values. It should be particularly noted that the calling to the data element protection catalogue is ^{required, or in other words} compelling. This means that in a system, in which the method according to the invention is implemented, ~~is such as to imply that~~ a user, who for instance wants to read a certain data element value in a given record in the first database, by his attempt at access to the data element value automatically ~~and compellingly~~ produces a system calling to the data element protection catalogue in the second database for collecting the protection attributes associated with the corresponding data element types. The continued processing procedure (reading of data element value) of the system is also controlled compellingly in accordance with the collected protection attribute/attributes applying to the corresponding data element types.

The term "data element protection catalogue" and the use thereof according to the invention must not be confused with the known term "active dictionary", which means that, in addition to an operative database, there is a special table indicating different definitions or choices for data element values in the operative database, for instance that a data element value "yellow" in terms of definition means a ^{colour} ~~colour~~ code which is within a numeric interval stated in such a reference table.

Preferably, the processing rules stated by the protection attributes are inaccessible to the user, and the read or collected protection attributes are preferably used merely internally by the system for controlling the processing. A given user, who, for instance, wants to read information stored in the database regarding a certain individual, thus need not at all be aware of the fact that certain protection attributes have been activated and resulted in certain, sensitive information for this individual being excluded from the information that is made available on e.g. a display. Each user-initiated measure aiming at processing of data element values thus

involves on the one hand a compelling calling to the data element protection catalogue and, on the other hand, a continued processing which is compellingly subjected to those processing rules that are stated by the protection attributes, and this may thus be accomplished without the user obtaining information on what rules control the processing at issue, and especially without the user having any possibility of having access to the rules.

By altering, adding and removing protection attributes in the data element protection catalogue, the person responsible for the system or an equivalent person may easily determine, for each individual data element type, the processing rules applying to data element values associated with the individual data element type and thus easily maintain a high and clear safety quality in the system.

According to the invention, it is thus the individual data element (data element type) and not the entire register that becomes the controlling unit for the way in which the ^{organization}~~organisation~~, operator etc. responsible for the system has determined the level of quality, responsibility and safety regarding the management of information.

To obtain a high level of protection, the data element protection catalogue is preferably encrypted so as to prevent ^{unauthorized}~~unauthorised~~ access thereto.

As preferred protection attributes, the present invention provides the following possibilities, which, however, are to be considered an incomplete, exemplifying list:

1. Statement of what "strength" or "level" (for instance none, 1, 2...) of encryption is to be used for storing the corresponding data element values in the database. Different data element values within one and the same record may thus be encrypted with mutually different strength.

2. Statement of what "strength" or "level" (for instance none, 1, 2,...) of encryption is to be used for the corresponding data element values if these are to be transmitted on a net.
- 5 3. Statement of program and/or versions of program that are authorised to be used for processing the corresponding data element values.
- 10 4. Statement of "owner" of the data element type. Different data element values within one and the same record can thus have different owners.
- 15 5. Statement of sorting-out rules for the corresponding data element values, for instance, statement of method and time for automatic removal of the corresponding data element values from the database.
- 20 6. Statement whether automatic logging is to be made when processing the corresponding data element values.

According to a specially preferred embodiment of the invention, the above-described PTY storing method is used for encryption of all data that is to be encrypted in both the database (i.e. the data element values) and the data element protection catalogue (i.e. the protection attributes). In the normal case where each record has a record identifier (corresponding to SID above), preferably also the record identifier is protected by means of PTY. Specifically, a floating alteration of the record identifiers in both the operative database and the data element protection catalogue can be made at desired intervals and at randomly selected times, in accordance with the above-described PTY principle. In the preferred embodiment, especially the encapsulated processor which is used for the PTY encryption can also be used for im-

plementation of the callings to the data element protection catalogue and the procedure for processing according to the collected protection attributes.

The invention will now be explained in more detail
5 with reference to the accompanying drawings, which schematically illustrate the inventive principle implemented in an exemplifying data system.

Brief Description of the Drawings

10 Fig. 1 (prior art) schematically shows the principle of storing of data information according to the PTY principle in W095/15628.

Fig. 2 (prior art) schematically shows the principle of producing floating storing identities according to the
15 PTY principle in W095/15628.

Fig. 3 schematically shows a computer system for implementing the method according to the invention.

Fig. 4 schematically shows the principle of data processing according to the invention with compelling
20 callings to a data element protection catalogue.

Fig. 5 shows an example of a display image for determining of protection attributes in the data element protection catalogue.

25 Description of the Preferred Embodiment

In the following, the designation IAM (which stands for Information Assets Manager) will be used for the components and applications which in the embodiment are essential to the implementation of the invention.

30 - Reference is first made to Fig. 3, which schematically illustrates a data managing system, in which the present invention is implemented and in which the following databases are included for storing information, in this example person-related information:

35 - An open database P-DB which contains generally accessible data, such as personal name, article

name, address etc. with the personal code number Pcn as plain text as record identifier;

- An operative database O-DB, which contains data that is to be protected. Encrypted identification, in this case an encrypted personal code number, is used as record identifier (= storage identity SID). O-DB is used by authorised users for processing of individual records, such as reading and update;
- An archive-database A-DB, which contains data transferred (sorted out) from the operative database O-DB and which is used for statistic questions, but not for questions directed to individual records. The transfer from O-DB to A-DB may take place in batches.
- A database IAM-DB, which is a database essential to the implementation of the invention. This database contains a data element protection catalogue with protection attributes for such data element types as are associated with data element values in records in the operative database O-DB. This database IAM-DB is preferably physically separated from the other O-DB and is inaccessible to the user. However, two or more sets of the data element protection catalogue may be available: on the one hand an original version to which only an authorised IAM operator has access and, on the other hand, a copy version which imports the data element protection catalogue from the original version and which may optionally be stored on the same file storage as the operative database O-DB. The two versions may be remote from each other, for instance be located in two different cities.

The data system in Fig. 3 further comprises a hardware component 10, a control module 20 (IAM-API), and a program module 30 (PTY-API). The function of these three components will now be described in more detail.

Hardware Component 10

The hardware component 10 acts as a distributed processor of its own in a computer. It has an encapsulation that makes it completely tamper-proof, which means that monitoring by so-called trace tools will not be possible.

The hardware component 10 can as an independent unit perform at least the following functions:

- Creating variable, reversible and non-reversible encrypting algorithms for the PTY encryption and providing these algorithms with the necessary variables;
- Initiating alterations of storage identities (SID) in stored data according to PTY, on the one hand data in O-DB and, on the other hand, data in the data element protection catalogue of IAM-DB;
- Storing user ^{authorisations} ~~authorisations~~ having access to records in O-DB; and
- Linking original identities OID to the correct record in O-DB.

Control Module 20 (IAM-API)

The control module ²⁰ controls the handling of the types of data protection that the system can supply.

The control module carries out the processing requested via API (Application Program Interface) programming interface.

Program Module 30 (PTY-API) 30

The program module (PTY-API) 30 handles the dialogue between the application 40 involved (including ACS) and the hardware component 10. This module may further log events and control sorting out/removal of data from the operative database O-DB.

Reference is now made to Fig. 4, which illustrates the same four databases (P-DB, O-DB, A-DB, IAM-DB) as in Fig. 3 and which schematically illustrates how the pro-

cessing of individual data elements are, according to the invention, controlled according to the rules that are stated by protection attributes in the data element protection catalogue, which is stored in the database
5 IAM-DB.

The data that is to be stored concerns in this example a certain individual and contains: (1) generally accessible data such as name and address, (2) identifying information, such as personal code number (Pcn), and (3)
10 descriptive information (DI). The generally accessible data name and address is stored together with personal code number (Pcn) in the open database P-DB, said storage being performable as plain text since this information is of the type that is generally accessible.

15 For storing the identifying information in combination with the descriptive information DI, the following steps will, however, be made, in which the following designations are used to describe encrypting and decrypting algorithms. Generally speaking, the encrypting and
20 decrypting algorithms can be described as follows:

$$F_{\text{Type}}(\text{Random number}, \text{Input data}) = \text{Results}$$

wherein:

F designates a function.

25 Type indicates the type of function as follows:
FKIR = Non-reversible encrypting algorithm
FKR = Reversible encrypting algorithm
FDKR = Decrypting algorithm

30 Random number
 represents one or more constants and/or
 variables included in the function F.

Input data

35 are the data to be encrypted or decrypted, and

Results indicate a unique function value for a given
 function

Step 1 Division of the information

Identifying information is separated from descriptive information;

5 Step 2 Preparation of storage identity SID:

10 An original identity OID is selected based on the identifying information. OID is here selected to be equal to the personal code number Pcn of the individual. OID is encrypted by means of a non-reversible encrypting algorithm ALG1, prepared randomly by the hardware component 10, to an update identity UID as follows:

15 ALG1: $F_{KIR}(\text{Random number}, \text{OID}) = \text{UID}$

ALG1 is such that attempts at decryption of UID to OID result in a great number of identities, which makes it impossible to link a specific UID to the corresponding OID.

20 Then UID is encrypted by means of a reversible algorithm ALG2, which is also produced at random by the hardware component 10, for generating a storage identity SID as follows:

25 ALG2: $F_{KR}(\text{Random number}, \text{UID}) = \text{SID}$

30 ALG2 is such that there exists a corresponding decrypting algorithm ALG3, by means of which SID can be decrypted in order to recreate UID.

35 The storage identity SID is used, as described in step 4 above, as encrypted record identifier when storing encrypted data element values DV in the operative database O-DB.

14

Step 3 Production of encrypted data element values DV:

The descriptive information DI associated with the original identity OID is converted into one or more encrypted data element values DV linked to a data element type DT each.

The encryption takes place as described below with a reversible encryption function F_{KR} , which like the algorithms ALG1 and ALG2 above is also produced at random by the hardware component 10. The invention is distinguished by a compelling calling here being sent to the data element protection catalogue in the database IAM-DB for automatic collection of the protection attribute which is linked to the data element type at issue and which indicates "strength" or degree with which the encryption of the descriptive data is to be performed so as to generate the data element value DV.

The table, which in Fig. 4 is shown below the database IAM-DB, ^{symbolises} ~~symbolises~~ an exemplifying content of the data element protection catalogue, here designated ^{DPC} ~~DC~~. As an example, it may here be assumed that the protection function $Func1$ corresponds to "degree of encryption". If the descriptive information DI at issue is to be stored as a data element value ^{DV} ~~associated~~ with the specific data element type DT1 in the data element protection catalogue, the protection attribute "5" registered in the data element protection catalogue is collected automatically in this case. The descriptive information DI at issue will thus, automatically and compellingly, be encrypted with the strength "5" for generating an encrypted data element value DV as follows:

15

15

F_{KR} (Random number, DI) = encrypted data element value DV

5 For storing a less sensitive data element, for instance a data element of the data element type DT3, the compelling calling to the data element protection catalogue in IAM-DB would instead have resulted in the protection attribute "no" being collected, in which case no encryption would have been made on the descriptive data at issue, which then could be stored as plain text in the operative database ODB.

15 Step 4 Storing of records in the operative database O-DB:
The encrypted storage identity SID according to step 2 in combination with the corresponding encrypted data element value or data element values DV according step 3 are stored as a record in the operative database O-DB.

20

As appears from the foregoing, a stored information record P has the following general appearance:

TO160

	Descript. information in the form of encrypted data element values			
Storage identity (SID)	DV1	DV2	DV3	DV4

25 The original identity OID is encrypted according to the PTY principle in two steps, of which the first is non-reversible and the second is reversible. Thus, it is impossible to store the descriptive information DI along with a storage identity SID that never can be linked to the original identity OID, as well as to create "floating", i.e. which change over time, storage identities SID while retaining the possibility of locating, for a specific original identity OID, the associated descriptive information DI stored.

30

16

The descriptive data DI is stored in accordance with protection attributes linked to each individual data element. This results in a still higher level of protection and a high degree of flexibility as to the setting up of rules, and continuous adaptation thereof, of how sensitive data is allowed to be used and can be used, down to the data element level.

✓ To increase the level of protection still more, the data element protection catalogue ^{DPC} ~~DE~~ is preferably stored in IAM-DB in encrypted form in accordance with the PTY principle, in which case for instance the data element types correspond to the above storage identity and the protection attributes correspond to the descriptive information or data element values above, as schematically illustrated in Fig. 4. This efficiently prevents every attempt at circumventing the data element protection by ^{unauthorised} ~~unauthorised~~ access and interpretation of the content of the data element protection catalogue.

✓ In the illustrated embodiment, "PTY can thus have the following functions:

- Protecting the original identity OID in encrypted form (SID) on the operative database O-DB (as is known from said WO95/15628),
- Protecting information in IAM-DB, particularly the protection attributes of the data element protection catalogue and the associated record identifier, and
- Protecting descriptive information DI in the form of encrypted data element values DV for the data element types that have the corresponding protection activated in the data element protection catalogue, and in accordance with the corresponding protection attributes.

Functionality Protection

35 In the above embodiment of the procedure for inputting data in the operative database O-DB, only "degree of encryption " has so far been discussed as data element

17

protection attribute in the data element protection catalogue DC. However, this is only one example among a number of possible protection attributes in the data element protection catalogue, which normally offers a plurality of protection attitudes for each data element. Preferred protection attributes have been indicated above in the general description.

A particularly interesting protection attribute is "protected programs". The use of this data element protection attribute means that the data system may offer a new type of protection, which is here called "functionality protection" and which means that only accepted or certified programs are allowed to be used and can be used in the system in the processing of data. It should be noted that this type of protection is still, according to the invention, on the data element level.

Now assume for the purpose of illustration that Func2 in the data element protection catalogue ^{DPC} ~~DC~~ in Fig. 4 corresponds to this protection attribute and that data elements of the data element type DT1 and DT2, respectively, are only allowed to ^{be} ~~processed~~ with the accepted applications or programs P1 and P2, respectively. ^{Unauthorized} ~~Unauthorized~~ handling of the corresponding data elements by means of, for instance, a different program P3, or a modified version P1' of P1, should be prevented. As protection attribute in the data element protection catalogue, data identifying P1 and P2 is therefore stored. In a preferred example, an cryptographic check sum P1* and P2*, respectively, is created, in a manner known per se, based on every accepted program P1 and P2, respectively. These check sums may be considered to constitute a unique fingerprint of the respective accepted programs, and these fingerprints can be stored as protection attributes in the data element protection catalogue as illustrated schematically in Fig. 4. It should however be noted that such check sums for accepted programs can optionally be stored in a data element protection catalogue of their

own for registering of accepted programs, separately from the data element protection catalogue with protection attributes for encryption strength.

If the last-mentioned type of protection "protected programs" is used, it should also be noted that the system, in connection with a user-initiated measure aiming at processing of a given data element, for instance inputting a new data element value in a certain record, need not carry out a complete examination of all programs accepted in the system. If, for instance, the user tries to use a program P3 for inputting in the operative database O-DB a new data element value, a compelling calling is sent to the data element protection catalogue in connection with the corresponding data element type, for instance DT1. The associated protection attribute P1* is then collected from the data element protection catalogue, which means that such a data element value is only allowed to be stored by means of the program P1. The attempt at registering the data element value by means of the program P3 would therefore fail.

By periodic use of the above-described functionality protection, it is possible to reveal and/or prevent that an ^{unauthorized} ~~unauthorised~~ person (for instance a "hacker") breaks into the system by means of a non-accepted program and modifies and/or adds descriptive data in such a manner that the descriptive data will then be identifying for the record. The data element values are thus not allowed to become identifying in the operative database O-DB.

30

Traceability/logging

"Logging" or "traceability" is another type of protection which according to the invention can be linked to a data element type in the data element protection catalogue. If this protection is activated for a certain data element type, each processing of the corresponding data element values in the operative database O-DB will auto-

atically and compellingly result in relevant information on the processing ("user", "date", "record", "user program" etc.) being logged in a suitable manner, so that based on the log, it is possible to investigate afterwards who has processed the data element values at issue, when, by means of which program etc.

Reading of Data from the Operative Database O-DB

In connection with a user-initiated measure aiming at reading/altering data element values in the stored records in the operative database O-DB, the following steps are carried out, which specifically also comprise a compelling calling to the data element protection catalogue and "unpacking" of the data which is controlled automatically and compellingly by collected protection attributes.

Step 1 The record is identified by producing the storage identity SID at issue based on the original identity OID, (Pcn) that is associated with the data element value DV which is to be read, as follows

$$FKR(FKIR(OID)) = SID$$

Step 2 When the record has been found by means of SID, the encrypted data element value DV (i.e. the encrypted descriptive data that is to be read) is decrypted as follows by means of a decrypting algorithm FDKR:

$$FDKR(DV) = \text{descriptive data (plain text)}$$

The carrying out of this decryption of the data element value, however, requires that the encryption-controlling protection attribute of the data element is first collected by the system from the data element protection catalogue DC, i.e. the

attribute indicating with which strength or at which level the data element value DV stored in O-DB has been encrypted. Like in the above procedure for inputting of data in O-DB, also when
5 reading, a compelling calling thus is sent to the data element protection catalogue DC for collecting information which is necessary for carrying out the processing, in this case the unpacking.

10 It will be appreciated that such a compelling calling to the data element protection catalogue DC, when making an attempt at reading, may result in the attempt failing, wholly or partly, for several reasons, depending on the protection
15 attribute at issue, which is linked to the data element value/values that is/are to be read. For instance, the attempt at reading may be interrupted owing to the user trying to use a non-accepted program and/or not being ^{authorized} ~~authorised~~ to
20 read the term involved.

If the data element protection catalogue is encrypted, the decoding key can be stored in a storage position separate from the first and the second database.

25 Fig. 5 shows an example of a user interface in the form of a dialogue box, by means of which a person responsible for IAM, i.e. a person responsible for security, may read and/or alter the protection attributes stated in the data element protection catalogue. In the
30 Example in Fig. 5, the data element types "Housing allowance" and "Social allowance" have both been provided with protection attributes concerning encryption, sorting out, logging and owner. Moreover, registration of ^{authorized} ~~authorised~~
C users and protected programs linked to the data element
35 type "Social allowance" has taken place in submenus.